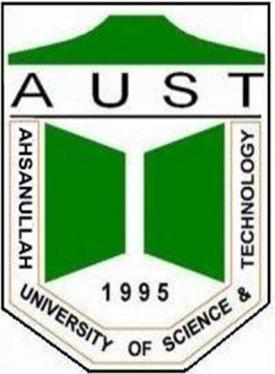
**Ahsanullah University of Science and Technology**



Department of Computer Science and Engineering

Program: Bachelor of Science in Computer Science and Engineering

Course No: CSE 4108

Course Title: Artificial Intelligence Lab

Project Report On: Laptop Price Prediction

Date of Submission: 05/09/2022

Submitted to:

Mr. Md. Siam Ansary

Lecturer, Department of CSE, AUST.

Ms. Tamanna Tabassum

Lecturer, Department of CSE, AUST.

Submitted by,

Name:  Md. Mominul Islam Shizan

Student ID: 18.02.04.117

Group: B2

Name:  Afridi Rahman Bondhon

Student ID: 18.02.04.128

Group: B2

**Introduction:**

Based on many practical characteristics of laptops from different companies, we will provide pricing predictions. Customers will find it easier to choose the right laptop for them and the best laptop configuration for their needs while keeping their budget in mind.

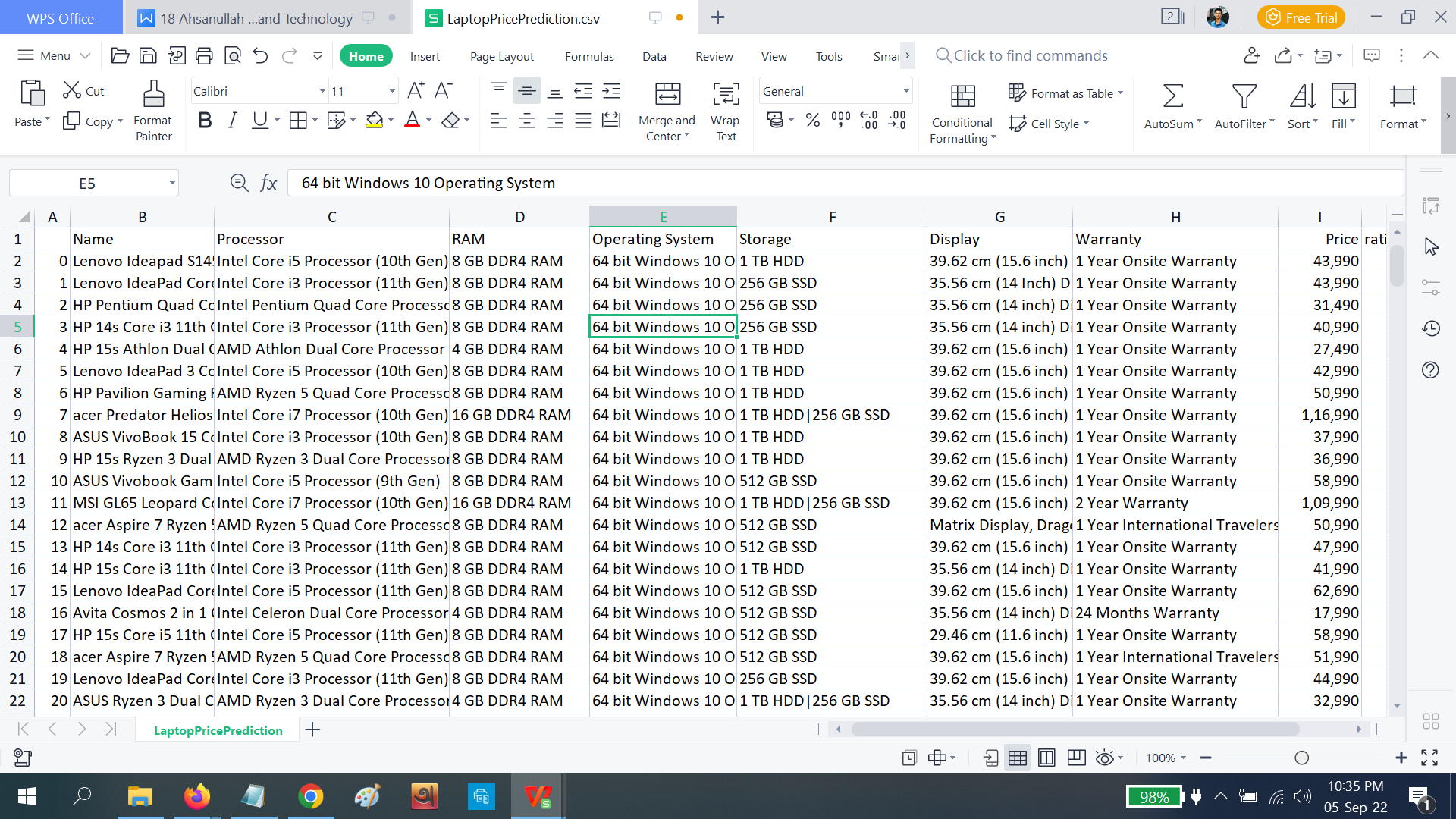
We created a dataset mergeing ours and an existing dataset and then split it into 67/33 train-test and subsequently tested the accuracy of our findings after applying various machine learning methods.

**Dataset:**

We have collected data from startech, ryanscomputers and some online shops. And Also from Kaggle.

Number of rows: 550

Number of attributes: 10



**ML Model:**

* **Linear Regression**

Linear Regression is a machine learning algorithm based on supervised learning. It performs a regression task. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting.

* **KNearestNeighbours**

KNN regression is a non-parametric method that, in an intuitive manner, approximates the association between independent variables and the continuous outcome by averaging the observations in the same neighborhood. The size of the neighborhood needs to be set by the analyst or can be chosen using cross-validation to select the size that minimizes the mean-squared error.

* **Random Forest**

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML.

* **Gradient Boosting**

Gradient Boosting is a popular boosting algorithm. In gradient boosting, each predictor corrects its predecessor’s error. In contrast to Adaboost, the weights of the training instances are not tweaked, instead, each predictor is trained using the residual errors of predecessor as labels.

* **Extra Trees**

Extra Trees is an ensemble machine learning algorithm that combines the predictions from many decision trees.

It is related to the widely used random forest algorithm. It can often achieve as-good or better performance than the random forest algorithm, although it uses a simpler algorithm to construct the decision trees used as members of the ensemble.

* **Voting regressor**

A voting regressor is an ensemble meta-estimator that fits several base regressors, each on the whole dataset. Then it averages the individual predictions to form a final prediction.

**Comparison of Performance Scores:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Name of  Algorithm | Mean  Absolute  Error | Mean  Squared  Error(%) | Root  Mean  Square  Error | Root  Mean  Square  Error(%) | R2  Score |
| Linear Regression | **21239.4564766006** | **26.10419891517** | **30500.47817498129** | **37.48639001967** | **0.5242991208967516** |
| K Nearest Neighbor | **27520.9483516483** | **33.8244206435879** | **40623.31615120926** | **49.92779012841** | **0.1561380531365369** |
| Random Forest | **23801.3508869701** | **29.2528765360767** | **34109.20994950024** | **41.92167546011** | **0.4050725904325450** |
| Gradient Boosting | **22402.3809354683** | **27.5334827309391** | **31848.29707719222** | **39.14291700112** | **0.4813276591312762** |
| Extra Tree | **26387.4856593406** | **32.4313465969178** | **37675.233216962544** | **46.30447032176** | **0.2741739408374113** |
| Voting Regressor | **21567.9283303520** | **26.5079048488855** | **30996.95774380058** | **38.09658460897** | **0.5086863837708089** |

Linear Regression performed better than other models

**Discussion:**

If we train our model with a larger datasets, we can achieve even better results.